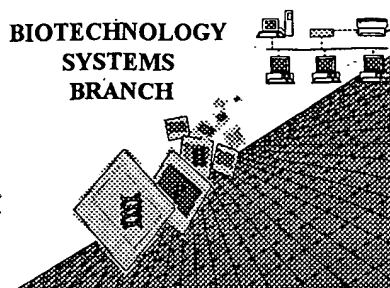


0400

**BEST AVAILABLE COPY**

**RAW SEQUENCE LISTING**  
**ERROR REPORT**



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/015,989  
Source: OIPK  
Date Processed by STIC: 1/3/2002

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE: SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom, including:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)

2. U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202

3. Hand Carry directly to:

U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7<sup>th</sup> Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202

Or

U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

# Raw Sequence Listing Error Summary

## ERROR DETECTED

## SUGGESTED CORRECTION

SERIAL NUMBER: 10/015,989

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1    Wrapped Nucleics  
    Wrapped Aminos      The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2    Invalid Line Length      The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3    Misaligned Amino  
    Numbering      The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4    Non-ASCII      The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5    Variable Length      Sequence(s)        contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6    PatentIn 2.0  
    "bug"      A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s)       . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7    Skipped Sequences  
    (OLD RULES)      Sequence(s)        missing. If intentional, please insert the following lines for each skipped sequence:  
                          (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
                          (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
                          (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
                          This sequence is intentionally skipped  
  
                          Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8    Skipped Sequences  
    (NEW RULES)      Sequence(s)        missing. If intentional, please insert the following lines for each skipped sequence.  
                          <210> sequence id number  
                          <400> sequence id number  
                          000
- 9    Use of n's or Xaa's  
    (NEW RULES)      Use of n's and/or Xaa's have been detected in the Sequence Listing.  
                          Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
                          In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10    Invalid <213>  
    Response      Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11    Use of <220>      Sequence(s)        missing the <220> "Feature" and associated numeric identifiers and responses.  
                          Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
                          (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12    PatentIn 2.0  
    "bug"      Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13    Misuse of n      n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

OICE

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/015,989

DATE: 01/03/2002  
TIME: 15:41:56

Input Set : A:\ES.txt  
Output Set: N:\CRF3\01032002\J015989.raw

*see item 4 on  
Error Summary  
Sheet*

**Does Not Comply** *app 1-12*  
**Corrected Diskette Needed**

7 <110> APPLICANT: DARROW, ANDREW  
10 QI, JENSON  
13 ANDRADE-GORDON, PATRICIA  
19 <120> TITLE OF INVENTION: ZYMOGEN ACTIVATION SYSTEM  
25 <130> FILE REFERENCE: ORT-1552  
31 <140> CURRENT APPLICATION NUMBER: US/10/015,989  
34 <141> CURRENT FILING DATE: 2001-12-10  
40 <160> NUMBER OF SEQ ID NOS: 60  
46 <170> SOFTWARE: PATENTIN VER. 2.0

## ERRORED SEQUENCES

52 <210> SEQ ID NO: 1  
55 <211> LENGTH: 361  
58 <212> TYPE: DNA  
61 <213> ORGANISM: ARTIFICIAL SEQUENCE  
67 <220> FEATURE:  
70 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE: FUSION GENE  
73 VECTORS.  
79 <400> SEQUENCE: 1  
E--> 82 gaattcacca ccatggacag caaagggttcg tgcagaaat cccgcctgct  
83 cctgctgctg 60  
E--> 86 gtggtgtcaa atctactctt gtgccagggt gtggtctccg actacaagga  
87 cgacgacgac 120  
E--> 90 gtggacgcgg ccgctcttgc tgcccccttt gatgatgatg acaagatcgt  
91 tgggggctat 180  
E--> 94 gctctagata gggccgctt ccctttagtg agggttaatg cttcgagcag  
95 acatgataag 240  
E--> 98 atacattgat gagtttggac aaaccacaac tagaatgcag tgaaaaaaat  
99 gctttatttg 300  
E--> 102 tgaaatttgt gatgctattg ctttatttgt aaccattata agctgcaata  
103 aacaagttga 360  
106 c  
112 <210> SEQ ID NO: 2  
115 <211> LENGTH: 301  
118 <212> TYPE: DNA  
121 <213> ORGANISM: ARTIFICIAL SEQUENCE  
127 <220> FEATURE:  
130 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE: FUSION GENE  
133 VECTORS.  
139 <400> SEQUENCE: 2  
E--> 142 gaattcacca tgaatccact cctgatcctt acctttgtgg cggccgctct  
143 tgetgcccc 60  
E--> 146 tttgatgatg atgacaagat cgttgggggc tattgtctag ataccctac  
147 gatgtgccg 120  
E--> 150 attagccta gggccgctt ccctttagtg agggttaatg cttcgagcag

*FYI: nucleotides must  
be in lower-case letters,  
when sequence listing is*

*in new  
sequence rules  
format*

*60 ← format (see  
120 ← error (item 1  
on  
Error  
Summary  
Sheet)*

361

*same format error*

DATE: 01/03/2002

TIME: 15:41:56

Output Set: N:\CRF3\01032002\J015989.raw

same

301

same

484

= same

## RAW SEQUENCE LISTING

DATE: 01/03/2002

PATENT APPLICATION: US/10/015,989

TIME: 15:41:56

Input Set : A:\ES.txt

Output Set: N:\CRF3\01032002\J015989.raw

E--> 286 gtgaaaaaaaa tgctttattt gtgaaatttg tgatgctatt gctttatttg  
 287 taaccattat 360  
 290 aagctgcaat aaacaagttg ac  
 296 <210> SEQ ID NO: 5  
 299 <211> LENGTH: 352  
 302 <212> TYPE: DNA  
 305 <213> ORGANISM: ARTIFICIAL SEQUENCE  
 311 <220> FEATURE:  
 314 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE: FUSION GENE  
 317 VECTORS.  
 323 <400> SEQUENCE: 5  
 E--> 326 gaattcacca ccatggcttt cctctggtc ctctcctgct gggccctcct  
 327 ggggtaccacc 60  
 E--> 330 ttcggctgcg ggggtccccga ctacaaggac gacgacgacg cggccgctct  
 331 tgetgcccc 120  
 E--> 334 tttgatgatg atgacaagat cggtgggggc tatgctctag acatcaccat  
 335 caccatcact 180  
 E--> 338 agcggccgct tccctttagt gagggttaat gcttcgagca gacatgataa  
 339 gatacattga 240  
 E--> 342 tgagtttggg caaacacaa ctagaatgca gtgaaaaaaaa tgctttattt  
 343 gtgaaatttg 300  
 E--> 346 tgatgctatt gctttatttg taaccattat aagctgcaat aaacaagttg  
 347 ac 352  
 353 <210> SEQ ID NO: 6  
 356 <211> LENGTH: 385  
 359 <212> TYPE: DNA  
 362 <213> ORGANISM: ARTIFICIAL SEQUENCE  
 368 <220> FEATURE:  
 371 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE: FUSION GENE  
 374 VECTORS.  
 380 <400> SEQUENCE: 6  
 E--> 383 gaattcacca ccatggcttt cctctggtc ctctcctgct gggccctcct  
 384 ggggtaccacc 60  
 E--> 387 ttcggctgcg ggggtccccga ctacaaggac gacgacgacg cggccgctct  
 388 tgetgcccc 120  
 E--> 391 tttgatgatg atgacaagat cggtgggggc tatgctctag ataccctac  
 392 gatgtgccc 180  
 E--> 395 attacgccgc tagacatcac catcaccatc actagcggcc gcttcccttt  
 396 agtgagggtt 240  
 E--> 399 aatgcttgcg gcagacatga taagatacat tgatgagttt ggacaaacca  
 400 caactagaat 300  
 E--> 403 gcagtgaaaa aaatgcttta tttgtgaaat ttgtgatgct attgctttat  
 404 ttgtaaccat 360  
 407 tataagctgc aataaacaag ttgac  
 413 <210> SEQ ID NO: 7  
 416 <211> LENGTH: 1169  
 419 <212> TYPE: DNA  
 422 <213> ORGANISM: ARTIFICIAL SEQUENCE  
 428 <220> FEATURE:

382

385

## RAW SEQUENCE LISTING

DATE: 01/03/2002

PATENT APPLICATION: US/10/015,989

TIME: 15:41:56

Input Set : A:\ES.txt

Output Set: N:\CRF3\01032002\J015989.raw

431 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE: FUSION GENE  
 434 WITH HOMO SAPIEN SERINE PROTEASE CATALYTIC DOMAIN  
 440 <400> SEQUENCE: 7

E--> 443 gaattcacca ccatggacag caaagggttcg tcgcagaaat cccgcctgct  
 444 cctgctgctg 60

E--> 447 gtgggtgtcaa atctactctt gtgccagggt gtgggtctccg actacaagga  
 448 cgacgacgac 120

E--> 451 gtggacgcgg ccgctcttgc tgcccccttt gatgatgatg acaagatcgt  
 452 tgggggctat 180

E--> 455 gctctagagg ccggtcagtg gccctggcag gtcagcatca cctatgaagg  
 456 cgtccatgtg 240

E--> 459 tgtgggtggct ctctcgtgtc tgagcagtggt gtgctgtcag ctgctcactg  
 460 cttccccagc 300

E--> 463 gagcaccaca aggaagccta tgagggtcaag ctggggggccc accagctaga  
 464 ctctactctc 360

E--> 467 gaggacgcca aggtcagcac cctgaaggac atcatcccc accccagcta  
 468 cctccaggag 420

E--> 471 ggctcccagg gcgacattgc actcctccaa ctacgacagc ccatacctt  
 472 ctcccgtctc 480

E--> 475 atccggccca tctgcctccc tgcagccaac gcctccttcc ccaacggcct  
 476 ccactgcact 540

E--> 479 gtcactgggt ggggtcatgt ggccccctca gtgagcctcc tgacgcccaa  
 480 gccactgcag 600

E--> 483 caactcgagg tgccctctgat cagtcgtgag acgtgtaact gcctgtacaa  
 484 catcgacgcc 660

E--> 487 aagcctgagg agccgcactt tgtccaagag gacatggtgt gtgctggcta  
 488 tgtggagggg 720

E--> 491 ggcaaggacg cctgccagggt tgactctggg ggccccactct cctgccctgt  
 492 ggagggtctc 780

E--> 495 tgggtacctga cgggcattgt gagctgggga gatgcctgtg gggccccgaa  
 496 caggcctggt 840

E--> 499 gtgtacactc tggcctccag ctatgcctcc tggatccaaa gcaagggtgac  
 500 agaactccag 900

E--> 503 cctcgtgtgg tgccccaaac ccaggagtcc cagcccgaca gcaacctctg  
 504 tggcagccac 960

E--> 507 ctggccttca gctctagaca tcaccatcac catcactagc ggccgcttcc  
 508 ctttagtgag 1020

E--> 511 ggttaatgct tcgagcagac atgataagat acattgatga gtttggacaa  
 512 accacaacta 1080

E--> 515 gaatgcagtg aaaaaaatgc tttatttggt aaatttggtga tgctattgct  
 516 ttatttgtaa 1140

519 ccattataag ctgcaataaa caagttgac

525 <210> SEQ ID NO: 8

528 <211> LENGTH: 1142

531 <212> TYPE: DNA

534 <213> ORGANISM: ARTIFICIAL SEQUENCE

540 <220> FEATURE:

543 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE: FUSION GENE  
 546 WITH HOMO SAPIEN SERINE PROTEASE CATALYTIC DOMAIN

*same*

1169

## RAW SEQUENCE LISTING

DATE: 01/03/2002

PATENT APPLICATION: US/10/015,989

TIME: 15:41:56

Input Set : A:\ES.txt

Output Set: N:\CRF3\01032002\J015989.raw

552 &lt;400&gt; SEQUENCE: 8

E--> 555 gaattcacca ccatggcttt cctctggctc ctctcctgct gggccctcct  
 556 ggggtaccacc 60  
 E--> 559 ttcggtgcg ggggtcccgga ctacaaggac gacgacgacg cggccgctct  
 560 tgcgtgcccc 120  
 E--> 563 tttgatgatg atgacaagat cgttgggggc tatgctctag aggccgggtca  
 564 gtggccctgg 180  
 E--> 567 caggtcagca tcacctatga aggcgtccat gtgtgtggtg gctctctcgt  
 568 gtctgagcag 240  
 E--> 571 tgggtgctgt cagctgctca ctgcttcccc agcgagcacc acaaggaagc  
 572 ctatgaggtc 300  
 E--> 575 aagctggggg cccaccagct agactcctac tccgaggacg ccaaggctcag  
 576 caccctgaag 360  
 E--> 579 gacatcatcc cccacccag ctacctccag gagggctccc agggcgacat  
 580 tgcactctc 420  
 E--> 583 caactcagca gacctatcac cttctcccgc tacatccggc ccacttgct  
 584 cctgcagcc 480  
 E--> 587 aacgcctcct tcccacagg cctccactgc actgtcactg gctgggggtca  
 588 tgtggcccc 540  
 E--> 591 tcagtgaacc tctgacgcc caagccactg cagcaactcg aggtgcctct  
 592 gatcagtcgt 600  
 E--> 595 gagacgtgta actgcctgta caacatcgac gccaaacctg aggagccgca  
 596 ctttgtccaa 660  
 E--> 599 gaggacatgg tgtgtgctgg ctatgtggag gggggcaagg acgcctgcca  
 600 gggtgactct 720  
 E--> 603 gggggccccc tctcctgccc tgtggagggt ctctgggtacc tgacgggcat  
 604 tgtgagctgg 780  
 E--> 607 ggagatgcct gtggggcccg caacaggcct ggtgtgtaca ctctggcctc  
 608 cagctatgcc 840  
 E--> 611 tcctggatcc aaagcaaggt gacagaactc cagcctcgtg tggtgcccca  
 612 aaccagag 900  
 E--> 615 tccagcccg acagcaacct ctgtggcagc cacctggcct tcagctctag  
 616 acatcaccat 960  
 E--> 619 caccatcact agcggccgct tccctttagt gagggttaat gcttcgagca  
 620 gacatgataa 1020  
 E--> 623 gatacattga tgagtttga caaaccacaa ctagaatgca gtgaaaaaaaa  
 624 tgctttattt 1080  
 E--> 627 gtgaaatttg tgatgctatt gctttatttg taaccattat aagctgcaat  
 628 aaacaagttg 1140  
 631 ac

1142

637 &lt;210&gt; SEQ ID NO: 9

640 &lt;211&gt; LENGTH: 1049

643 &lt;212&gt; TYPE: DNA

646 &lt;213&gt; ORGANISM: ARTIFICIAL SEQUENCE

652 &lt;220&gt; FEATURE:

655 &lt;223&gt; OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE: FUSION GENE

658 WITH HOMO SAPIEN SERINE PROTEASE CATALYTIC DOMAIN

664 &lt;400&gt; SEQUENCE: 9

E--&gt; 667 gaattcacca ccatggacag caaaggttcg tcgcagaaat cccgcctgct

## RAW SEQUENCE LISTING

DATE: 01/03/2002

PATENT APPLICATION: US/10/015,989

TIME: 15:41:56

Input Set : A:\ES.txt

Output Set: N:\CRF3\01032002\J015989.raw

668 cctgctgctg 60  
 E--> 671 gtggtgtcaa atctactctt gtgccagggt gtggtctccg actacaagga  
 672 cgacgacgac 120  
 E--> 675 gtggacgagg ccgctcttgc tgcccccttt gatgatgatg acaagatcgt  
 676 tggggggtac 180  
 E--> 679 aactgtctag aaccccatc gcagccttgg caggcggcct tgttcaggg  
 680 ccagcaacta 240  
 E--> 683 ctctgtggcg gtgtccttgt aggtggcaac tgggtcctta cagctgccca  
 684 ctgtaaaaaa 300  
 E--> 687 ccgaaataca cagtacgcct gggagaccac agcctacaga ataaagatgg  
 688 ccagagcaa 360  
 E--> 691 gaaatacctg tggttcagtc catcccacac ccctgctaca acagcagcga  
 692 tgtggaggac 420  
 E--> 695 cacaaccatg atctgatgct tcttcaactg cgtgaccagg catccctggg  
 696 gtccaaagtg 480  
 E--> 699 aagcccatca gcctggcaga tcattgcacc cagcctggcc agaagtgcac  
 700 cgtctcaggc 540  
 E--> 703 tggggcactg tcaccagtcc ccgagagaat tttctgaca ctctcaactg  
 704 tgcagaagta 600  
 E--> 707 aaaatcttcc ccagaagaa gtgtgaggat gcttaccggg ggcagatcac  
 708 agatggcatg 660  
 E--> 711 gtctgtgcag gcagcagcaa aggggctgac acgtgccagg gcgattctgg  
 712 agggccctg 720  
 E--> 715 gtgtgtgatg gtgcactcca gggcatcaca tctgggggt cagaccctg  
 716 tgggaggtcc 780  
 E--> 719 gacaaacctg gcgtctatac caacatctgc cgctacctgg actggatcaa  
 720 gaagatcata 840  
 E--> 723 ggcagcaagg gctctagaca tcaccatcac catcactagg ggcgcttcc  
 724 ctttagtgag 900  
 E--> 727 ggttaatgct tcgagcagac atgataagat acattgatga gtttggacaa  
 728 accacaacta 960  
 E--> 731 gaattcagtg aaaaaaatgc tttatttgtg aaatttgtga tgctattgct  
 732 ttatttgtaa 1020  
 735 ccattataag ctgcaataaa caagttgac  
 741 <210> SEQ ID NO: 10  
 744 <211> LENGTH: 1052  
 747 <212> TYPE: DNA  
 750 <213> ORGANISM: ARTIFICIAL SEQUENCE  
 756 <220> FEATURE:  
 759 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE: FUSION GENE  
 762 WITH HOMO SAPIEN SERINE PROTEASE CATALYTIC DOMAIN  
 768 <400> SEQUENCE: 10  
 E--> 771 gaattcacca ccatggacag caaagggttcg tcgcagaaat cccgcctgct  
 772 cctgctgctg 60  
 E--> 775 gtggtgtcaa atctactctt gtgccagggt gtggtctccg actacaagga  
 776 cgacgacgac 120  
 E--> 779 gtggacgagg ccgctcttgc tgcccccttt gatgatgatg acaagatcgt  
 780 tggggggtac 180  
 E--> 783 aactgtctag aaaagcactc ccagccctgg caggcagccc tgttcgagaa

1049

## RAW SEQUENCE LISTING

DATE: 01/03/2002

PATENT APPLICATION: US/10/015,989

TIME: 15:41:56

Input Set : A:\ES.txt

Output Set: N:\CRF3\01032002\J015989.raw

784 gacgcggcta 240  
 E--> 787 ctctgtgggg cgacgtcat cgccccaga tggctcctga cagcagccca  
 788 ctgcctcaag 300  
 E--> 791 ccccgctaca tagttcacct ggggcagcac aacctccaga aggaggagg  
 792 ctgtgagcag 360  
 E--> 795 acccggaacag ccactgagtc cttccccac cccggcttca acaacagcct  
 796 ccccaacaaa 420  
 E--> 799 gaccaccgca atgacatcat gctgggtgaag atggcatcgc cagtctccat  
 800 cacctgggct 480  
 E--> 803 gtgcgacccc tcacctctc ctcacgtctgt gtcactgctg gcaccagctg  
 804 cctcatttcc 540  
 E--> 807 ggctggggca gcacgtccag cccccagtta cgctgcctc acaccttgcg  
 808 atgcgccaac 600  
 E--> 811 atcaccatca ttgagcacca gaagtgtgag aacgcctacc ccggcaacat  
 812 cacagacacc 660  
 E--> 815 atgggtgtgtg ccagcgtgca ggaagggggc aaggactcct gccaggggtga  
 816 ctccgggggc 720  
 E--> 819 cctctgggtct gtaaccagtc tcttcaaggc attatctcct ggggccagga  
 820 tccgtgtgcg 780  
 E--> 823 atcaccgaa agcctgggtgt ctacacgaaa gtctgcaaat atgtggactg  
 824 gatccaggag 840  
 E--> 827 acgatgaaga acaattctag acatcaccat caccatcact agcggccgct  
 828 tccctttagt 900  
 E--> 831 gagggttaat gcttcgagca gacatgataa gatacattga tgagtttgga  
 832 caaaccacaa 960  
 E--> 835 ctagaatgca gtgaaaaaaaa tgctttattt gtgaaatttg tgatgctatt  
 836 gctttatttg 1020  
 839 taaccattat aagctgcaat aaacaagttg ac

1052

1067 <210> SEQ ID NO: 12  
 1070 <211> LENGTH: 319  
 1073 <212> TYPE: PRT  
 1076 <213> ORGANISM: ARTIFICIAL SEQUENCE  
 1082 <220> FEATURE:  
 1085 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE: FUSION GENE  
 1088 WITH HOMO SAPIEN SERINE PROTEASE CATALYTIC DOMAIN  
 1094 <400> SEQUENCE: 12

1097 MET ALA PHE LEU TRP LEU LEU SER CYS TRP ALA LEU LEU GLY THR THR  
 1100 1 5 10 15  
 1106 PHE GLY CYS GLY VAL PRO ASP TYR LYS ASP ASP ASP ASP ALA ALA ALA  
 1109 20 25 30  
 1115 LEU ALA ALA PRO PHE ASP ASP ASP ASP LYS ILE VAL GLY GLY TYR ALA  
 1118 35 40 45  
 1124 LEU GLU ALA GLY GLN TRP PRO TRP GLN VAL SER ILE THR TYR GLU GLY  
 1127 50 55 60  
 1133 VAL HIS VAL CYS GLY GLY SER LEU VAL SER GLU GLN TRP VAL LEU SER  
 1136 65 70 75 80  
 1142 ALA ALA HIS CYS PHE PRO SER GLU HIS HIS LYS GLU ALA TYR GLU VAL  
 1145 85 90 95  
 1151 LYS LEU GLY ALA HIS GLN LEU ASP SER TYR SER GLU ASP ALA LYS VAL

see  
P. 8, too

Per sequence  
 Rules (1.822)  
 only the first  
 letter of  
 amino acid is  
 in upper-case.  
 e.g. Met Ala

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/015,989

DATE: 01/03/2002

TIME: 15:41:56

Input Set : A:\ES.txt

Output Set: N:\CRF3\01032002\J015989.raw

```

1154          100          105          110
1160 SER THR LEU LYS ASP ILE ILE PRO HIS PRO SER TYR LEU GLN GLU GLY
1163          115          120          125
1169 SER GLN GLY ASP ILE ALA LEU LEU GLN LEU SER ARG PRO ILE THR PHE
1172          130          135          140
1178 SER ARG TYR ILE ARG PRO ILE CYS LEU PRO ALA ALA ASN ALA SER PHE
1181 145          150          155          160
1187 PRO ASN GLY LEU HIS CYS THR VAL THR GLY TRP GLY HIS VAL ALA PRO
1190          165          170          175
1196 SER VAL SER LEU LEU THR PRO LYS PRO LEU GLN GLN LEU GLU VAL PRO
1199          180          185          190
1205 LEU ILE SER ARG GLU THR CYS ASN CYS LEU TYR ASN ILE ASP ALA LYS
1208          195          200          205
1214 PRO GLU GLU PRO HIS PHE VAL GLN GLU ASP MET VAL CYS ALA GLY TYR
1217          210          215          220
1223 VAL GLU GLY GLY LYS ASP ALA CYS GLN GLY ASP SER GLY GLY PRO LEU
1226 225          230          235          240
1232 SER CYS PRO VAL GLU GLY LEU TRP TYR LEU THR GLY ILE VAL SER TRP
1235          245          250          255
1241 GLY ASP ALA CYS GLY ALA ARG ASN ARG PRO GLY VAL TYR THR LEU
E--> 1242 ALA
E--> 1245          260          265          270
1251 SER SER TYR ALA SER TRP ILE GLN SER LYS VAL THR GLU LEU GLN PRO
E--> 1254          275          280          285
1260 ARG VAL VAL PRO GLN THR GLN GLU SER GLN PRO ASP SER ASN LEU CYS
E--> 1263          290          295          300
1269 GLY SER HIS LEU ALA PHE SER SER ARG HIS HIS HIS HIS HIS HIS
E--> 1272 305          310          315
2409 <210> SEQ ID NO: 35
2412 <211> LENGTH: 55
2415 <212> TYPE: DNA
2418 <213> ORGANISM: ARTIFICIAL SEQUENCE
2424 <220> FEATURE:
2427 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE:
2430 OLIGONUCLEOTIDE
2436 <400> SEQUENCE: 35
E--> 2439 aattcaccac catggtttc ctctggctcc tctcctgctg ggccctcctg
2440 ggtac 55
2446 <210> SEQ ID NO: 36
2449 <211> LENGTH: 47
2452 <212> TYPE: DNA
2455 <213> ORGANISM: ARTIFICIAL SEQUENCE
2461 <220> FEATURE:
2464 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE:
2467 OLIGONUCLEOTIDE
2473 <400> SEQUENCE: 36
E--> 2476 ccaggagggc ccagcaggag aggagccaga ggaaagccat ggtgggtg
2477 47
2483 <210> SEQ ID NO: 37

```

*convert  
last two letters of  
amino acids  
to lower-case  
letters*

*move up*

*move up - see  
item 1  
on Error  
summary  
sheet*

*same  
error*

## RAW SEQUENCE LISTING

DATE: 01/03/2002

PATENT APPLICATION: US/10/015,989

TIME: 15:41:56

Input Set : A:\ES.txt

Output Set: N:\CRF3\01032002\J015989.raw

2486 <211> LENGTH: 45  
 2489 <212> TYPE: DNA  
 2492 <213> ORGANISM: ARTIFICIAL SEQUENCE  
 2498 <220> FEATURE:  
 2501 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE:  
 2504 OLIGONUCLEOTIDE  
 2510 <400> SEQUENCE: 37  
 E--> 2513 **caccttcggc tgcgggggtcc ccgactacaa ggacgacgac gacgc** *same*  
 2514 45  
 2520 <210> SEQ ID NO: 38  
 2523 <211> LENGTH: 53  
 2526 <212> TYPE: DNA  
 2529 <213> ORGANISM: ARTIFICIAL SEQUENCE  
 2535 <220> FEATURE:  
 2538 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE:  
 2541 OLIGONUCLEOTIDE  
 2547 <400> SEQUENCE: 38  
 E--> 2550 **ggccgcgtcg tgcgtcgtcct tgtagtcggg gaccccgag ccgaaggtgg** *same*  
 2551 tac 53  
 2629 <210> SEQ ID NO: 41  
 2632 <211> LENGTH: 55  
 2635 <212> TYPE: DNA  
 2638 <213> ORGANISM: ARTIFICIAL SEQUENCE  
 2644 <220> FEATURE:  
 2647 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE:  
 2650 OLIGONUCLEOTIDE  
 2656 <400> SEQUENCE: 41  
 E--> 2659 **ggccgctctt gctgccccct ttgatgatga tgacaagatc gttgggggct** *same*  
 2660 atgct 55  
 2666 <210> SEQ ID NO: 42  
 2669 <211> LENGTH: 55  
 2672 <212> TYPE: DNA  
 2675 <213> ORGANISM: ARTIFICIAL SEQUENCE  
 2681 <220> FEATURE:  
 2684 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE:  
 2687 OLIGONUCLEOTIDE  
 2693 <400> SEQUENCE: 42  
 E--> 2696 **ctagagcata gcccacaacg atcttgtcat catcatcaaa gggggcagca** *same*  
 2697 agagc 55  
 2703 <210> SEQ ID NO: 43  
 2706 <211> LENGTH: 55  
 2709 <212> TYPE: DNA  
 2712 <213> ORGANISM: ARTIFICIAL SEQUENCE  
 2718 <220> FEATURE:  
 2721 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE:  
 2724 OLIGONUCLEOTIDE  
 2730 <400> SEQUENCE: 43  
 E--> 2733 **ggccgctctt gctgccccct ttgatgatga tgacaagatc gttgggggct** *same*  
 2734 attgt 55

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/015,989

DATE: 01/03/2002

TIME: 15:41:56

Input Set : A:\ES.txt

Output Set: N:\CRF3\01032002\J015989.raw

2740 <210> SEQ ID NO: 44  
 2743 <211> LENGTH: 55  
 2746 <212> TYPE: DNA  
 2749 <213> ORGANISM: ARTIFICIAL SEQUENCE  
 2755 <220> FEATURE:  
 2758 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE:  
 2761 OLIGONUCLEOTIDE  
 2767 <400> SEQUENCE: 44  
 E--> 2770 ctagacaata gcccccaacg atcttgtcat catcatcaaa gggggcagca *same*  
 2771 agagc 55  
 2777 <210> SEQ ID NO: 45  
 2780 <211> LENGTH: 52  
 2783 <212> TYPE: DNA  
 2786 <213> ORGANISM: ARTIFICIAL SEQUENCE  
 2792 <220> FEATURE:  
 2795 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE:  
 2798 OLIGONUCLEOTIDE  
 2804 <400> SEQUENCE: 45  
 E--> 2807 ggccgctctt gctgccccct ttatcgaggg ggcattgtg gagggtctgg *same*  
 2808 at 52  
 2814 <210> SEQ ID NO: 46  
 2817 <211> LENGTH: 52  
 2820 <212> TYPE: DNA  
 2823 <213> ORGANISM: ARTIFICIAL SEQUENCE  
 2829 <220> FEATURE:  
 2832 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE:  
 2835 OLIGONUCLEOTIDE  
 2841 <400> SEQUENCE: 46  
 E--> 2844 ctagatccga gccctccaca atgcgcccct cgataaaggg ggcagcaaga *same*  
 2845 gc 52  
 3280 <210> SEQ ID NO: 54  
 3283 <211> LENGTH: 284  
 3286 <212> TYPE: PRT  
 3289 <213> ORGANISM: ARTIFICIAL SEQUENCE  
 3295 <220> FEATURE:  
 3298 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE: HUMAN MH2  
 3301 PROTEASE IN PFEK ZYMOGEN VECTOR  
 3307 <400> SEQUENCE: 54  
 3310 MET ASP SER LYS GLY SER SER GLN LYS SER ARG LEU LEU LEU LEU LEU  
 3313 1 5 10 15  
 3319 VAL VAL SER ASN LEU LEU LEU CYS GLN GLY VAL VAL SER ASP TYR LYS  
 3322 20 25 30  
 3328 ASP ASP ASP ASP VAL ASP ALA ALA ALA LEU ALA ALA PRO PHE ASP ASP  
 3331 35 40 45  
 3337 ASP ASP LYS ILE VAL GLY GLY TYR ASN CYS LEU GLU PRO HIS SER GLN  
 3340 50 55 60  
 3346 PRO TRP GLN ALA ALA LEU VAL MET GLU ASN GLU LEU PHE CYS SER GLY  
 3349 65 70 75 80  
 3355 VAL LEU VAL HIS PRO GLN TRP VAL LEU SER ALA ALA HIS CYS PHE GLN

*please  
 edit  
 amino acid  
 letters*

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/015,989

DATE: 01/03/2002

TIME: 15:41:56

Input Set : A:\ES.txt

Output Set: N:\CRF3\01032002\J015989.raw

```

3358          85          90          95
3364 ASN SER TYR THR ILE GLY LEU GLY LEU HIS SER LEU GLU ALA ASP GLN
3367          100          105          110
3373 GLU PRO GLY SER GLN MET VAL GLU ALA SER LEU SER VAL ARG HIS PRO
3376          115          120          125
3382 GLU TYR ASN ARG PRO LEU LEU ALA ASN ASP LEU MET LEU ILE LYS LEU
3385          130          135          140
3391 ASP GLU SER VAL SER GLU SER ASP THR ILE ARG SER ILE SER ILE ALA
3394 145          150          155          160
3400 SER GLN CYS PRO THR ALA GLY ASN SER CYS LEU VAL SER GLY TRP GLY
3403          165          170          175
3409 LEU LEU ALA ASN GLY ARG MET PRO THR VAL LEU GLN CYS VAL ASN
E--> 3410 VAL
E--> 3413          180          185          190
3419 SER VAL VAL SER GLU GLU VAL CYS SER LYS LEU TYR ASP PRO LEU TYR
E--> 3422          195          200          205
3428 HIS PRO SER MET PHE CYS ALA GLY GLY GLY HIS ASP GLN LYS ASP SER
E--> 3431          210          215          220
3437 CYS ASN GLY ASP SER GLY GLY PRO LEU ILE CYS ASN GLY TYR LEU GLN
E--> 3440 225          230          235          240
3446 GLY LEU VAL SER PHE GLY LYS ALA PRO CYS GLY GLN VAL GLY VAL PRO
E--> 3449          245          250          255
3455 GLY VAL TYR THR ASN LEU CYS LYS PHE THR GLU TRP ILE GLU LYS THR
E--> 3458          260          265          270
3464 VAL GLN ALA SER SER ARG HIS HIS HIS HIS HIS HIS
E--> 3467          275          280
3608 <210> SEQ ID NO: 59
3611 <211> LENGTH: 1103
3614 <212> TYPE: DNA
3617 <213> ORGANISM: ARTIFICIAL SEQUENCE
3623 <220> FEATURE:
3626 <223> OTHER INFORMATION: DESCRIPTION OF ARTIFICIAL SEQUENCE: NUCLEIC ACID
3629     SEQUENCE OF HUMAN PROTEASE F IN CFK2 ZYMOGEN
3632     VECTOR
3638 <400> SEQUENCE: 59
E--> 3641 gaattcacca ccattgcttt cctctggctc ctctcctgct gggccctcct
3642 gggtaccacc 60
E--> 3645 ttctggctgcg gggtccccga ctacaaggac gacgacgacg cggccgctct
3646 tgctgcccc 120
E--> 3649 tttgatgatg atgacaagat cgttgggggc tatgctctag aactcgggcg
3650 ttggcctgtg 180
E--> 3653 caggggagcc tgcgcctgtg ggattccac gtatgaggag tgagcctgct
3654 cagccaccgc 240
E--> 3657 tgggcactca cggcggcgca ctgctttgaa acctatagtg accttagtga
3658 tccctccggg 300
E--> 3661 tggatggtcc agtttgcca gctgacttcc atgccatcct tctggagcct
3662 gcaggcctac 360
E--> 3665 tacaaccgtt acttcgtatc gaatatctat ctgagccctc gctacctggg
3666 gaattcaccc 420

```

*mod up**format  
env**see item 1  
on Env summary  
sheet*

## RAW SEQUENCE LISTING

DATE: 01/03/2002

PATENT APPLICATION: US/10/015,989

TIME: 15:41:56

Input Set : A:\ES.txt

Output Set: N:\CRF3\01032002\J015989.raw

E--> 3669 tatgacattg ccttggtgaa gctgtctgca cctgtcacct acactaaaca  
 3670 catccagccc 480  
 E--> 3673 atctgtctcc aggcctccac atttgagttt gagaaccgga cagactgctg  
 3674 ggtgactggc 540  
 E--> 3677 tgggggtaca tcaaagagga tgaggcactg ccatctcccc acaccctcca  
 3678 ggaagttcag 600  
 E--> 3681 gtgcgccatca taaacaactc tatgtgcaac cacctcttcc tcaagtacag  
 3682 ttcccgcaag 660  
 E--> 3685 gacatctttg gagacatggt ttgtgctggc aatgcccaag gcgggaagga  
 3686 tgctgtcttc 720  
 E--> 3689 ggtgactcag gtggaccctt ggctgtaac aagaatggac tgtggtatca  
 3690 gattggagtc 780  
 E--> 3693 gtgagctggg gagtgggctg tggctggccc aatcgggccg gtgtctacac  
 3694 caatatcagc 840  
 E--> 3697 caccactttg agtggatcca gaagctgatg gccagagtg gcattgtcca  
 3698 gccagacccc 900  
 E--> 3701 tcttgggtcta gacatcacca tcaccatcac tagcgggcgc ttccctttag  
 3702 tgagggttaa 960  
 E--> 3705 tgcttcgagc agacatgata agatacattg atgagtttgg acaaaccaca  
 3706 actagaatgc 1020  
 E--> 3709 agtgaaaaaa atgctttatt tgtgaaattt gtgatgctat tgctttattt  
 3710 gtaaccatta 1080  
 3713 taagctgcaa taaacaagtt gac

1103

*same  
error*



Review the Sequence Listing to insure a corresponding explanation is presented in the <220> to <223> fields of each sequence using n or Xaa.

## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/015,989

DATE: 01/03/2002

TIME: 15:41:57

Input Set : A:\ES.txt

Output Set: N:\CRF3\01032002\J015989.raw

L:31 M:270 C: Current Application Number differs, Replaced Application Number  
L:34 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:82 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:1  
L:82 M:112 C: (48) String data converted to lower case,  
M:112 Repeated in SeqNo=1  
M:254 Repeated in SeqNo=1  
L:142 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:2  
M:112 Repeated in SeqNo=2  
M:254 Repeated in SeqNo=2  
L:198 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:3  
M:112 Repeated in SeqNo=3  
M:254 Repeated in SeqNo=3  
L:266 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:4  
M:112 Repeated in SeqNo=4  
M:254 Repeated in SeqNo=4  
L:326 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:5  
M:112 Repeated in SeqNo=5  
M:254 Repeated in SeqNo=5  
L:383 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:6  
M:112 Repeated in SeqNo=6  
M:254 Repeated in SeqNo=6  
L:443 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:7  
M:112 Repeated in SeqNo=7  
M:254 Repeated in SeqNo=7  
L:555 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:8  
M:112 Repeated in SeqNo=8  
M:254 Repeated in SeqNo=8  
L:667 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:9  
M:112 Repeated in SeqNo=9  
M:254 Repeated in SeqNo=9  
L:771 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:10  
M:112 Repeated in SeqNo=10  
M:254 Repeated in SeqNo=10  
L:1242 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:12  
M:332 Repeated in SeqNo=12  
L:1719 M:112 C: (48) String data converted to lower case,  
L:1755 M:112 C: (48) String data converted to lower case,  
L:1791 M:112 C: (48) String data converted to lower case,  
L:1827 M:112 C: (48) String data converted to lower case,  
L:1863 M:112 C: (48) String data converted to lower case,  
L:1899 M:112 C: (48) String data converted to lower case,  
L:1935 M:112 C: (48) String data converted to lower case,  
L:1971 M:112 C: (48) String data converted to lower case,  
L:2007 M:112 C: (48) String data converted to lower case,  
L:2043 M:112 C: (48) String data converted to lower case,  
L:2079 M:112 C: (48) String data converted to lower case,  
L:2115 M:112 C: (48) String data converted to lower case,  
L:2151 M:112 C: (48) String data converted to lower case,

## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/015,989

DATE: 01/03/2002

TIME: 15:41:57

Input Set : A:\ES.txt

Output Set: N:\CRF3\01032002\J015989.raw

L:2187 M:112 C: (48) String data converted to lower case,  
L:2223 M:112 C: (48) String data converted to lower case,  
L:2259 M:112 C: (48) String data converted to lower case,  
L:2295 M:112 C: (48) String data converted to lower case,  
L:2331 M:112 C: (48) String data converted to lower case,  
L:2367 M:112 C: (48) String data converted to lower case,  
L:2403 M:112 C: (48) String data converted to lower case,  
L:2439 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:35  
M:112 Repeated in SeqNo=35  
L:2476 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:47 SEQ:36  
M:112 Repeated in SeqNo=36  
L:2513 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:45 SEQ:37  
M:112 Repeated in SeqNo=37  
L:2550 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:38  
M:112 Repeated in SeqNo=38  
L:2587 M:112 C: (48) String data converted to lower case,  
L:2623 M:112 C: (48) String data converted to lower case,  
L:2659 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:41  
M:112 Repeated in SeqNo=41  
L:2696 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:42  
M:112 Repeated in SeqNo=42  
L:2733 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:43  
M:112 Repeated in SeqNo=43  
L:2770 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:44  
M:112 Repeated in SeqNo=44  
L:2807 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:45  
M:112 Repeated in SeqNo=45  
L:2844 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:46  
M:112 Repeated in SeqNo=46  
L:2881 M:112 C: (48) String data converted to lower case,  
L:2917 M:112 C: (48) String data converted to lower case,  
L:2953 M:112 C: (48) String data converted to lower case,  
L:2989 M:112 C: (48) String data converted to lower case,  
L:3025 M:112 C: (48) String data converted to lower case,  
L:3061 M:112 C: (48) String data converted to lower case,  
L:3410 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:54  
M:332 Repeated in SeqNo=54  
L:3503 M:112 C: (48) String data converted to lower case,  
L:3641 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:59  
M:254 Repeated in SeqNo=59  
L:3752 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:60  
M:254 Repeated in SeqNo=60